

ESG Lab Review

ExaGrid with Veeam: Virtual Machine Backup without Compromise

Date: May 2013 **Author:** Kerry Dolan, Lab Analyst, and Vinny Choinski, Senior Lab Analyst

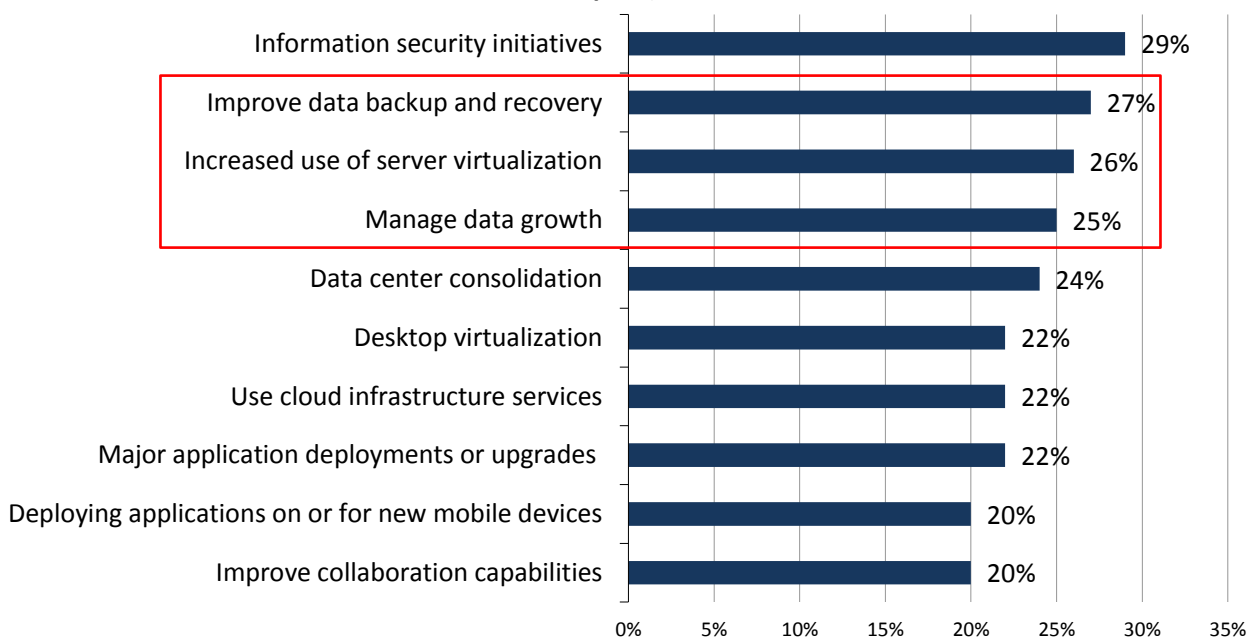
Abstract: This ESG Lab review documents hands-on testing of the ExaGrid backup appliance using Veeam Backup & Replication software, and demonstrates the ability to execute Veeam Instant VM Recovery alongside ExaGrid's capacity- and cost-saving deduplication.

The Challenges

A number of challenging tasks keep IT professionals up at night, and at times they converge to create a single bigger challenge. According to ESG research, improving data backup and recovery, increasing server virtualization, and managing data growth ranked very high among respondents to our most recent IT spending intentions survey (see Figure 1).¹ These three have occupied the top spots for several years now, and it's no wonder, as both virtualization and continually growing data volumes tend to make backup and recovery more complex.

Figure 1. Most Important IT Priorities

Which of the following would you consider to be your organization's most important IT priorities over the next 12 months? (Percent of respondents, N=540, ten responses accepted)



Source: Enterprise Strategy Group, 2013.

¹ Source: ESG Research Report, [2013 IT Spending Intentions Survey](#), January 2013.

The goal of ESG Lab reports is to educate IT professionals about data center technology products for companies of all types and sizes. ESG Lab reports are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objective is to go over some of the more valuable feature/functions of products, show how they can be used to solve real customer problems and identify any areas needing improvement. ESG Lab's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments. This ESG Lab report was sponsored by ExaGrid.

The flexibility of virtual machines (VMs) can cause serious managerial challenges. Deleting an entire VM unintentionally, including the applications installed on it, is easy to do. However, restoring VMs is not nearly as easy or fast when using the wrong approach. Backing up VMs provides assurance that in the case of an accidental deletion, failure, or disaster, you will be able to get the VMs back, so many organizations backup VMs with regularity. For recovery, users often have to wait hours or even days.

Instant Restore . . .

A recent technology innovation was born out of the intersection of disk-based backup and server virtualization. A few advanced backup and restore solutions enable VMs that are saved on disk to be “instantly recovered” by simply powering on VMs from the backup disk. This short-termed solution is designed to speed the user’s return to productivity so that time is not lost while IT finds and fixes the problem.

. . . Collides with Deduplication

However, a challenge for many vendors has arisen due to a conflict with another technology that has become crucial for efficiency and cost-savings as organizations work to manage ongoing data growth: deduplication. The capacity and TCO savings available from this technology (particularly for backups that continually store redundant data) cannot be overstated, and as a result it is rapidly becoming the norm for data protection. In the case of a disaster, IT can restore data from a deduplicated backup, but the time required often exceeds today’s recovery time objectives (RTOs).

So what about instant restore? Unfortunately, while instant restore works well with plain disk storage, it is not so effective with deduplicated backup targets. The rehydration process required to recover target-based deduplicated backups involves reassembling the data from small blocks and extensive hash tables. In addition, starting up a VM adds the strain of extremely random storage access, different from the sequential access typical of conventional full restores. These characteristics make the “instant” part of recovery virtually impossible, unless deduplication is optimized for disk-based backup.

This leaves IT organizations with a dilemma when protecting VMs: They must choose between gaining back productive time with instant restore, and minimizing storage capacity and cost with deduplication. Deduplication is crucial to reducing costs for virtual machines in particular, as VM backups often contain redundant copies of operating systems and applications. For many, instead of choosing based on the priority of the benefit, they implement multiple storage solutions for different data sets, driving up the costs of both equipment and management. For others, it means choosing one benefit and forgoing the other.

The Solution: ExaGrid Backup Appliance plus Veeam Backup & Replication

Together, [ExaGrid](#) and [Veeam](#) offer a solution that enables customers to enjoy the benefits of *both* deduplication and instant recovery when protecting virtual machines in VMware vSphere and Microsoft Hyper-V environments. In addition, customers benefit from the combination of Veeam source-side deduplication and ExaGrid post-process deduplication for maximum capacity reduction.

ExaGrid

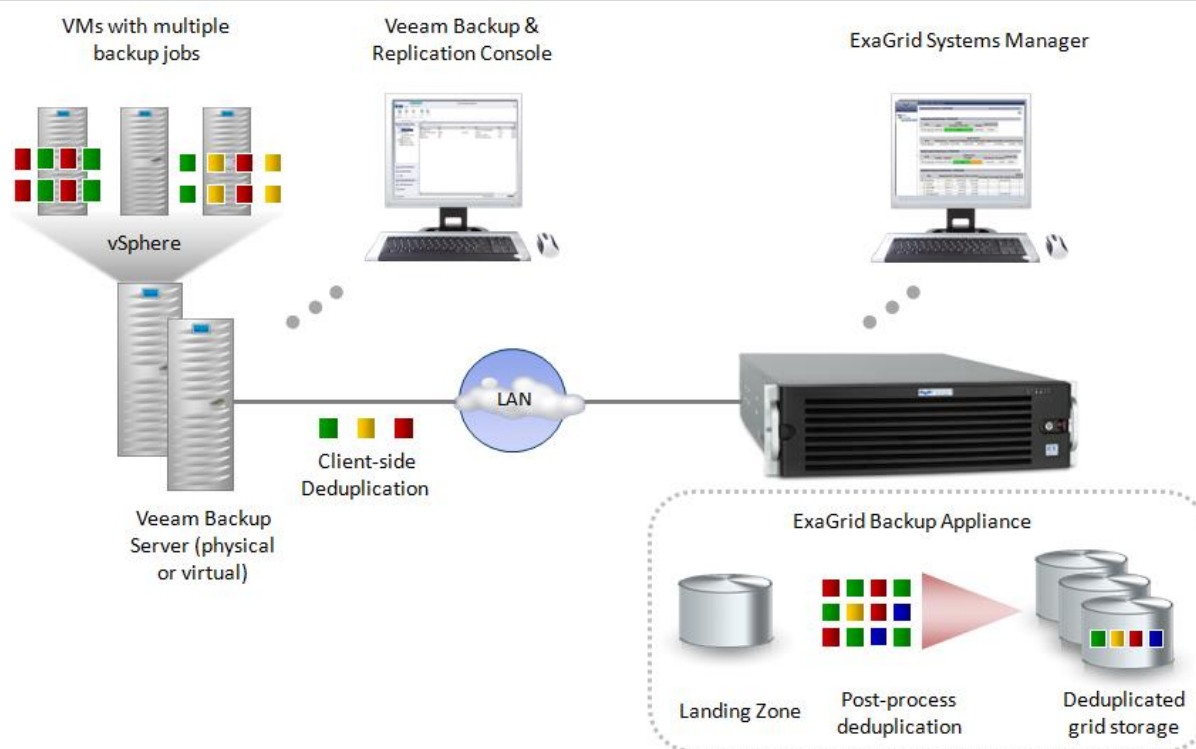
ExaGrid offers a series of high-performance, scalable disk backup appliances with built-in, post-process deduplication. Each ExaGrid appliance is a turnkey backup solution that works seamlessly with industry-leading backup software, and is built from components that include Intel quad-core Xeon processors, enterprise SATA drives in a RAID 6 configuration, and ExaGrid software. Appliances are simple to set up, speeding time to value, and protected data can be replicated directly to a remote ExaGrid appliance, offsite disk, or tape. A range of models includes up to 32TB of raw capacity per unit; these can be combined in grids of up to ten appliances for a total 320 TB of raw capacity capable of handling 130TB of primary data plus retention. This yields the equivalent of more than 2 PB of logical capacity that can be managed from a single console. As appliances are added to a grid, they each add CPU, bandwidth, memory, and disk capacity for linear scalability of capacity and performance.

Patented zone-level deduplication occurs after backups have completed (post-process, not inline during the backup process) so that backups are executed as quickly as possible. By storing only unique bytes across backups, ExaGrid provides capacity savings of 10:1 to 50:1 over time.

Landing Zone

A unique capability of ExaGrid appliances that enables instant restore is the “landing zone,” a high-speed cache where backups land before deduplication, and where a non-deduplicated copy of the most recent backup is held separately from the deduplicated backend (see Figure 2). Backups can be instantly recovered from the landing zone in just minutes, eliminating the wait time for a full restore from millions of deduplicated data chunks.

Figure 2. ExaGrid Appliance with Veeam Backup & Replication



Veeam Backup & Replication

Veeam Backup & Replication was built specifically for virtual environments, and offers application-aware backup and recovery, granular restore, and replication for VMware vSphere and Microsoft Hyper-V virtual machines. It is easy to set up and uses no agents in VMs, eliminating the need to license, deploy, manage, and maintain them. Veeam can be deployed with any disk storage, and its distributed architecture enables backup proxy VMs to reside at any location while leveraging the Veeam backup server for scheduling and load balancing. It can scale to support any environment by adding proxies and/or backup servers. Veeam's backup solution delivers powerful features including built-in deduplication and compression to reduce LAN bandwidth and storage capacity requirements.

Veeam's patented vPower technology enables the application to offer several unique capabilities:

- **Instant VM Recovery** – Veeam was a pioneer in instant recovery technologies. By exposing the virtual machine files in a backup repository through a disk mount, Veeam enables a backed-up VM to be *powered on and used* without having to wait for it to be fully restored. Local disk space serves as a temporary write area, recording changes while the VM is running from the backup. The full VM migration is conducted in the background, after which the VM running from the backup disk can be migrated to production storage for continued operation.
- **SureBackup** – Automatic verification of the recoverability of backups. Many organizations don't test the recoverability of backups because it requires additional dedicated resources and time. Instead, those responsible for disaster recovery simply hope that if needed, their backups will be recoverable. SureBackup automates verification using available resources in the production or test environment. Using Veeam technology, the VM is run in an isolated environment directly from the disk-based backup. Veeam starts the VM, boots the OS, confirms that the VM is functioning properly, and reports back to the VM administrator that the VM(s) are indeed recoverable. Application-aware confirmation can be scripted to match corporate policies, and results can be used for DR audits.

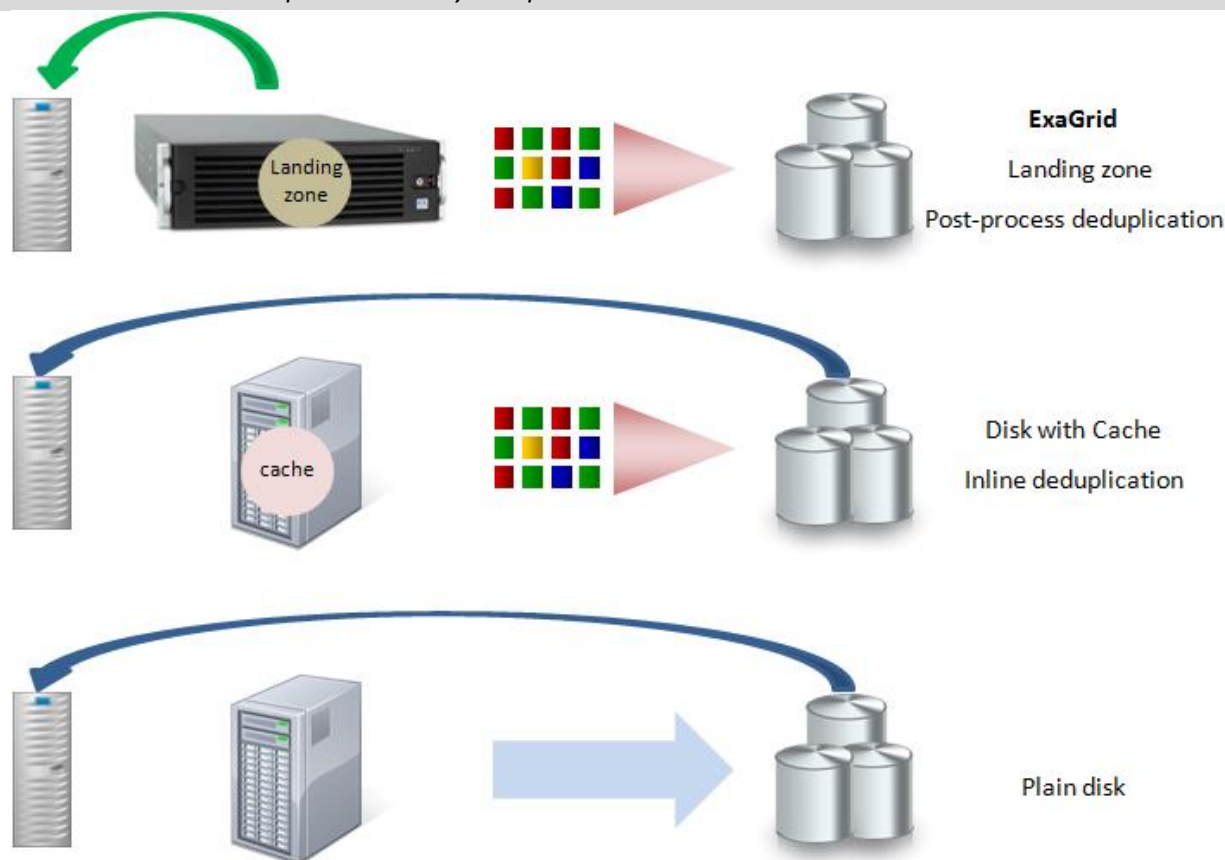
Other features include Universal Application-Item Recovery (U-AIR), which provides recovery of individual objects from virtualized applications; intelligent load balancing; a single view of centralized and remote office deployments from Veeam Enterprise Manager; and simple, per-socket licensing.

Instant VM Recovery

Instant recovery capabilities are closely tied to the type of storage used for backup. To provide some context, Figure 3 shows three different disk backup options and how they operate.

- At the bottom of the figure is plain disk storage, with no deduplication. The backups are sent directly to disk, stored as written by the backup application, and recovered directly from disk. Data can be instantly recovered from the disk repository.
- In the middle is an example of a disk backup target that includes a cache and in-line deduplication. Backups are deduplicated as the backup is being written to disk, trading some backup window time for the benefit of reduced disk capacity. The backup is not stored as written by the backup application, but instead is stored as deduplicated chunks. Standard restore is accomplished by rehydrating the data chunks from the disk in conjunction with a hash table. Instant recovery is not effective here as there is no non-deduplicated data. Instant recovery techniques can be executed, but may take hours (exceeding today's RTOs), and often fail.
- At the top of the figure is an ExaGrid disk target. The ExaGrid landing zone provides a high-speed cache where the last full (or synthetic full) backup and any incremental backups are stored as written by the backup application, while deduplication processing occurs on the appliance after the backup has finished. This enables the capacity savings of deduplication, while leaving a copy of the most recent backup intact and available for instant recovery. In an instant recovery scenario, VMs are powered on and run while accessing the selected recovery point stored in the ExaGrid landing zone.

Figure 3. Disk-based Backup and Recovery Comparison



ESG Lab Testing

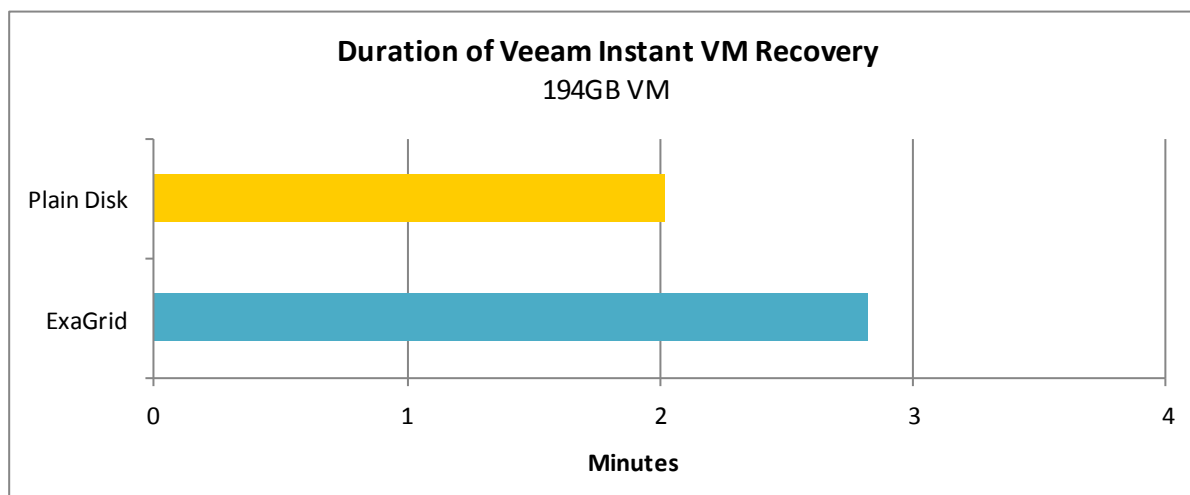
ESG Lab tested the Instant VM Recovery capability of the ExaGrid and Veeam solution, using both an ExaGrid appliance and plain disk as backup targets. Tests included duration of instant recovery, a comparison of end-user downtime for instant recovery versus full restore, and performance measurements of common workloads.

Testing was conducted using three Windows 2008 R2 virtual machines of different sizes (76GB, 129GB, and 194GB). Each VM consisted of a mix of data types, including OS/system files, an Exchange database, and common user files. Each VM was allocated 2 vCPUs, local storage, and 4GB RAM. Prior to ESG Lab’s onsite visit, a series of backups was completed on each VM using Veeam Backup & Replication version 6.5. In order to simulate a common production week, incremental backups were schedule and executed in three-hour intervals (to simulate daily incrementals), and one daily full backup was done (to simulate a weekly full). Most Veeam default settings were used, including Veeam server-side deduplication. The default exceptions were two storage settings, one for a LAN-based backup target, and the other to leverage Veeam’s “dedupe-friendly” compression (described later in this paper). Backups were completed for five days to simulate five weeks of backup workflow, and workload generators were used between full backups to simulate a weekly growth rate of 2%. All three VMs, plus a virtual Veeam Backup Server (which also functioned as a Proxy), were hosted on an Intel Xeon-based server with 16GB RAM, 43GB of SATA storage capacity, and two striped SSDs used for Veeam vPower temporary storage. VMware vSphere 5.0 provided the hypervisor.

Backups were directed to two targets: an ExaGrid EX10000E appliance (identified as *TCE10KS1OB1*) and an Intel Xeon-based Windows file server acting as a NAS disk target (identified as *WIN10G*). These targets were connected to the host using separate 10GbE LANs.

ESG Lab began testing by measuring the duration of Veeam Instant VM Recovery for both ExaGrid and plain disk. On the ExaGrid target, the latest full backup is held, non-deduplicated, in the ExaGrid landing zone. The Lab executed a Veeam Instant VM Recovery of the 194GB VM backup stored on the ExaGrid target by clicking on the **Instant VM Recovery** option from the Veeam Restore Wizard, and configuring for automatic power-on of the VM. Once the VM login screen appeared, we noted the elapsed time. The same procedure was executed for the 194GB VM backup located on the plain disk array (see Figure 4).

Figure 4. Instant VM Recovery Duration



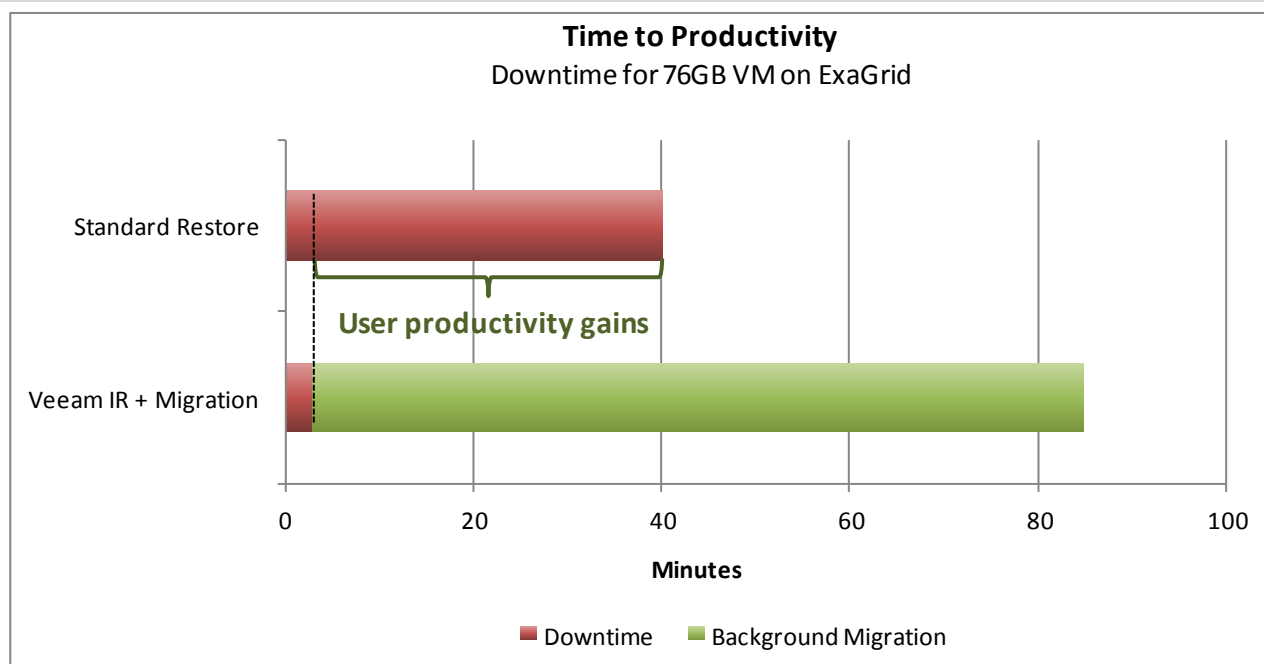
What the numbers mean:

- Instant VM Recovery from the plain disk target was completed in two minutes and 0.94 seconds (2:00.94).
- Instant VM Recovery from the ExaGrid target was completed in two minutes and 49.68 seconds (2:49.68).
- Both Instant VM Recoveries were completed in less than three minutes.

- Instant VM Recovery from the landing zone backup on ExaGrid took only 48.74 seconds longer than the plain disk backup.

Next, ESG Lab compared time to end-user productivity between a Veeam Instant VM Recovery and a full restore. The Lab executed a Veeam Instant VM Recovery of the 76GB VM stored on the ExaGrid target, with automatic VM power-on, and noted the elapsed time once the login screen appeared. Once the Instant VM Recovery was complete, the full migration to production storage ran in the background with the VM powered on, and the duration was recorded. Next, ESG Lab launched and timed a standard restore from the ExaGrid target (see Figure 5).

Figure 5. Time to Productivity



What the numbers mean:

- Instant VM Recovery of the 76GB VM from an ExaGrid backup took two minutes and 58 seconds (2:58).
- The background full migration of the 76GB VM after Instant VM Recovery took an additional 82 minutes, and the VM was online and usable during this process.
- Standard full restore of the same VM took 40 minutes.
- By using Veeam Instant VM Recovery, an additional 37.04 minutes of productive time was gained back over standard restore.
- With Instant VM Recovery, users were able to resume operations and productivity during the time it took to complete a full migration in the background.
- It should be noted that the time for full restore was 40 minutes for both ExaGrid and plain disk targets.

Next, ESG Lab measured the throughput and response times for workloads running on all three virtual machines. Using Iometer to simulate common workloads, ESG Lab ran measurements of I/O per second (IOPS) and response times for 8K OLTP running on the 76GB VM; Exchange 2010 running on the 129GB VM; and file server workloads running on the 194GB VM.

The Lab first ran all three workloads on the VMs using the local production storage to gather baseline data. Next, a Veeam Instant VM Recovery of the 129GB machine was launched on the ExaGrid appliance; the Exchange workload ran on the 129GB VM accessing the ExaGrid landing zone backup images, while the OLTP and file workloads ran on VMs operating from production storage. Finally, the workloads were run in the same configuration, but with the Exchange workload accessing the instantly recovered VM on plain disk. As shown in Figure 6, the IOPS results for each workload are color-coded: OLTP appears in blue, Exchange in red, and file server in green. The orange line indicates average response time for all workloads combined for each VM.

Figure 6. Throughput and Response Time

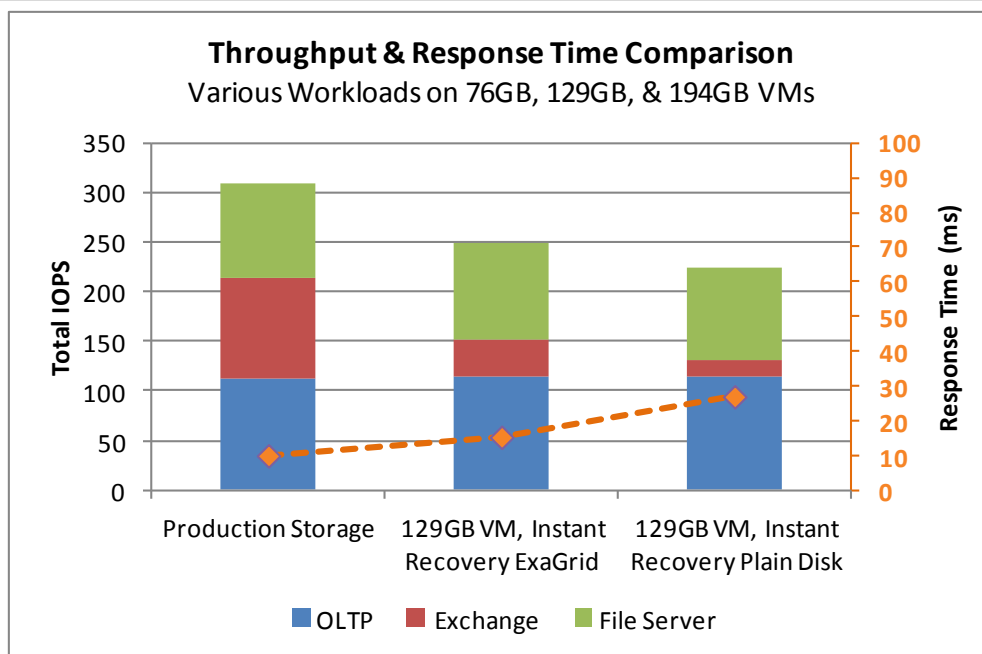


Table 1. Throughput and Response Time, Production vs. Instant Recovery (IR)

Workload (IOPS)	Production Storage	129GB VM, IR ExaGrid	129GB VM, IR plain disk
OLTP	112.41	114.75	114.67
Exchange 2010	100.99	35.99	16.23
File server	95.80	98.37	94.26
Aggregate	309.20	249.11	225.16
Response Time (ms)	Production Storage	129GB VM, IR ExaGrid	129GB VM, IR plain disk
Avg., aggregate workloads	9.97	15.31	26.97

What the numbers mean:

- Average response time was 15.31 ms for aggregated workloads on the ExaGrid target, which is only 5.3 ms longer than production storage.
- Average response time was almost 12 ms faster for the ExaGrid instant recovery than the plain disk instant recovery.
- Aggregate IOPS for all workloads remained high, with the ExaGrid IR generating 81% of the IOPS that production storage generated, and the plain disk IR generating 73%.
- Throughput was higher when Exchange was accessing the Exagrid target than when accessing plain disk.

Why This Matters

IT managers struggle to keep up with massive data growth created by the ease of spinning up virtual machines. Backing up all this data strains storage resources and complicates backup procedures. Deduplication can drastically reduce the amount of storage capacity required, improving TCO. Server virtualization also enables much higher application and data availability, to which end-users have rapidly become accustomed, resulting in more stringent SLA demands. Advanced backup applications can offer instant recovery to increase uptime, but instant recovery operates best with backups that have not been deduplicated on the backend. For any VM backup, IT must decide the priority, capacity savings, or high availability, and choose backup processes accordingly.

ESG Lab validated that a combination of the ExaGrid deduplicated backup appliance and Veeam Backup & Replication software can deliver both capacity savings and instant restore capability. In ESG Lab testing, the ExaGrid landing zone enabled the appliance to act like plain disk, enabling instant restore in less than three minutes. Testing also demonstrated the increase in productive time provided by the combined solution, saving 37 minutes over standard restore. Finally, ESG Lab validated that throughput and response times were very good when accessing an instantly recovered VM on ExaGrid, enabling users to function at a high level despite the degraded state. The ExaGrid target delivered better throughput and faster response time than the plain disk target.

Ease of Use and Efficiency

ESG Lab also tested the ease and efficiency of implementing the combined ExaGrid and Veeam solution, as well as the capacity savings provided by deduplication.

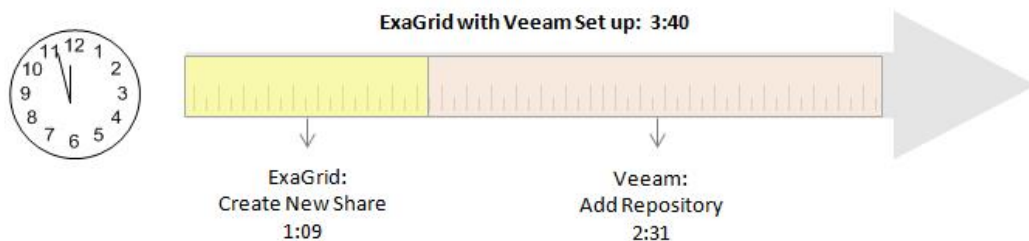
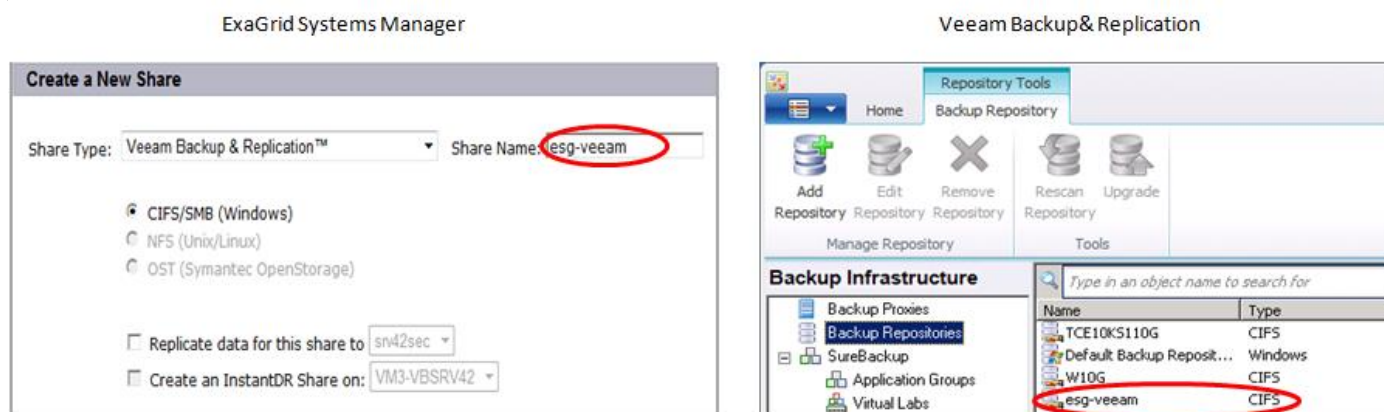
ESG Lab Tested

ESG Lab tested the ease of use and time needed to prepare the ExaGrid and Veeam solution for a new backup. This required using the web-based ExaGrid Systems Manager GUI to create a new Veeam share on the appliance, and using the Veeam Backup & Replication GUI to add a new backup repository.

From the ExaGrid Systems Manager GUI, the Lab selected **Create New Share** from the Manage Shares page. The Lab selected Veeam Backup & Replication from the Share Type drop-down menu, named the new share *esg-veeam*, and selected the radio button for CIFS/SMB (Windows). After IP-based security and a DNS host name were added, the Lab clicked **Create**, and the new share was ready. The total time for this task was one minute and nine seconds (1:09). Figure 7 shows key screens from the ExaGrid and Veeam GUIs, as well as a timeline.

Next, the Lab moved to the Veeam Backup & Replication GUI, selecting Backup Repository from the Repository Tools menu and clicking **Add Repository**. After typing in the name (*esg-veeam*) and description, ESG Lab selected the type (CIFS, shared folder), provided the IP address of the shared folder on the ExaGrid appliance, supplied credentials, and created the repository. At this point, the amount of available capacity was viewable. The final configuration selections included limiting Veeam concurrent jobs to ten, per Veeam best practice²; enabling the vPower NFS datastore using striped SSDs; and reviewing the configuration. The total elapsed time for Veeam set up was two minutes and 31 seconds (2:31). The combined time to set up both ExaGrid and Veeam was three minutes and 40 seconds (3:40); at this point, the Lab was ready to create backup jobs.

Figure 7. Set-up Time

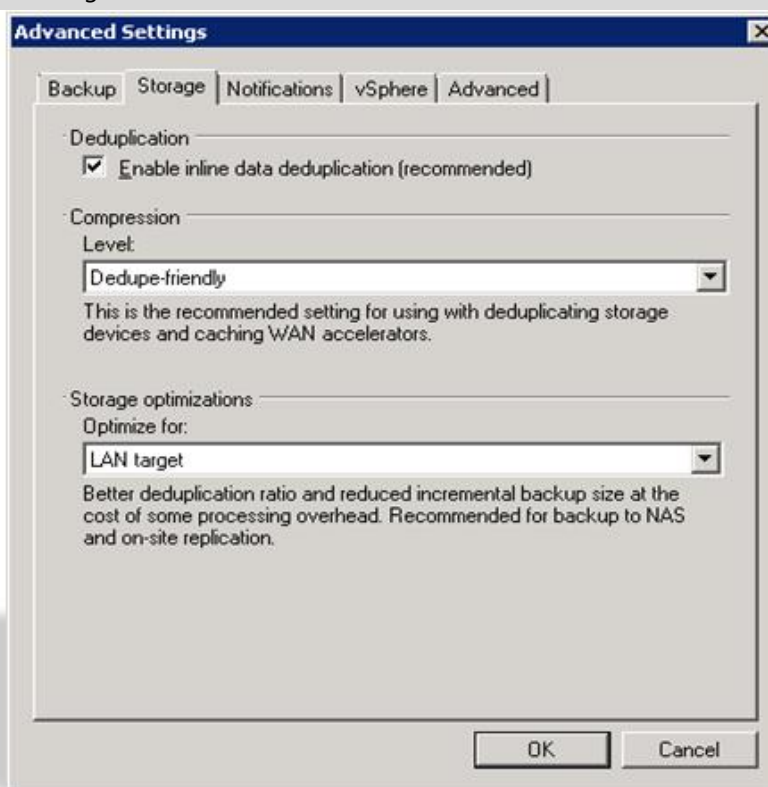


² While each ExaGrid appliance supports up to 20 concurrent backup jobs, Veeam suggests a limit of ten concurrent jobs per backup proxy.

Next, ESG Lab created a backup job for the three VMs using the *esg-veeam* repository on the ExaGrid appliance. Using the Veeam Backup & Replication GUI, the Lab named a new job **esg-all**, selected the 76GB, 129GB, and 194GB virtual machines from the Virtual Machine screen, and added them to the job.

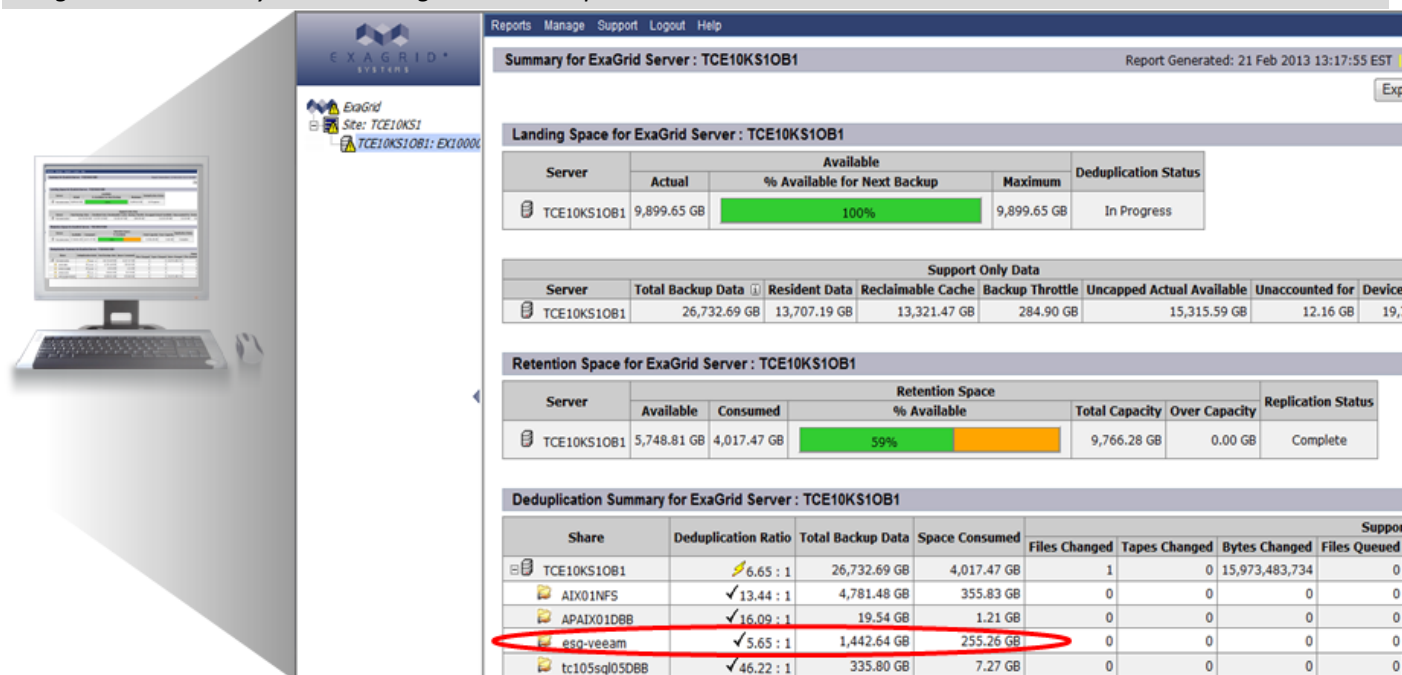
On the Storage screen, the Veeam backup proxy was automatically selected. For a backup repository, we selected the *esg-veeam* repository and configured a retention period. At the bottom of the Storage screen, ESG Lab clicked on the *Advanced* button to display additional configuration options (see Figure 8). From the Compression drop-down menu, **Dedupe-friendly** was selected. Veeam offers this setting to optimize backups for deduplicated disk targets; it provides source-side data reduction with a simpler algorithm, relieving some the overhead from Veeam. Finally, we selected **Optimize for LAN Target** from the storage optimizations menu, since ExaGrid presents a NAS-like interface to backup applications.

Figure 8. Veeam Advanced Storage Settings



Next, ESG Lab reviewed the ExaGrid deduplication statistics for the simulated five weeks of backups previously executed. From the ExaGrid Systems Manager, the Lab reviewed a capacity summary report for the appliance *TCE10KS10B1*. As Figure 9 shows, capacity details are displayed in tables for the landing space at the top and the total retention space in the middle. The bottom table shows a deduplication summary by share. As the table shows, the *esg-veeam* share backed up a total of 1.4 TB of data while consuming only 255GB of capacity, for a deduplication ratio of 5.65:1.

Figure 9. ExaGrid Systems Manager with Deduplication Details



Why This Matters

IT administrators juggle multiple tasks, so the easier and faster a task can be done, the better. The ability to set up backup jobs quickly and easily saves time and management effort—and ultimately money—especially as virtual environments grow. Deduplication of backups can enable huge savings in storage capacity needs, and as a result it is extremely popular. In recent ESG research, more than one-third (37%) of respondents reported currently using deduplication, with an additional 39% reporting plans to implement it within the next two years, for a total of 76%.³

ESG Lab validated the ease and efficiency of deploying the ExaGrid/Veeam solution; the total time to configure both products and be ready to create backup jobs was three minutes and 40 seconds (3:40). Veeam also includes advanced storage settings optimized for deduplicated backup targets. Finally, ESG Lab demonstrated the tremendous capacity savings available with the combined solution. From an initial data set of 2.1TB, Veeam performed client-side deduplication at a rate of 1.5:1, sending only 1.4TB to the ExaGrid target over a simulated five-week backup. Additional ExaGrid post-process deduplication at 5.6:1 further reduced that 1.4TB, resulting in only 255GB of storage consumed, and a combined deduplication ratio of 8.4:1. (It should be noted that testing was conducted with Veeam backup jobs using individual VMs, which does not produce typical deduplication rates. In most environments, jobs contain multiple VMs and therefore more redundancy.)

³ Source: ESG Research Report, [Trends in Data Protection Modernization](#), August 2012.

The Bigger Truth

Despite a preponderance of fancy new and hyped technologies, IT organizations must spend much of their time focused on day-to-day blocking and tackling. It's really a heroic effort to ensure that whatever new implementations come along, business users maintain high productivity, with as much access to applications and data as IT can provide. They do all this regardless of upgrades, maintenance, failures, compliance, security, and a myriad of other requests and requirements, and all within a context of cost containment. For, while some macroeconomic conditions have improved greatly since the 2008 global financial crisis, ESG research indicates that uncertainty around the "fiscal cliff" and other concerns has caused cost reduction initiatives to edge back into first place among items influencing IT spending.⁴

Unfortunately, these two needs—productivity and cost containment—can be opposing forces. Business managers must regularly balance improvements to business processes with the costs to implement them, without really knowing which will deliver the maximum value.

For many, this classic conundrum is represented in the choice between backup applications that can do instant recovery of virtual machines, and backups that provide deduplication to shrink storage capacity and cost. Instant recovery can be a boon to productivity, keeping virtual machine failures from hindering productivity for more than a few minutes. They are designed to be used like a "donut" spare tire—you don't expect to run on it for the long term, but if you can keep the car moving on four wheels while you find a gas station to fix the flat, your progress is not interrupted. Instant restore capabilities like Veeam Instant VM Recovery offer a similar benefit: Users can gain access to their VMs and return to productive operations quickly, while IT troubleshoots in the background. At the same time, deduplication has demonstrated a tremendous impact to storage capacity and cost, with deduplication rates for backup solutions such as ExaGrid shrinking storage needs by 95% and more.

ESG Lab validated that ExaGrid appliances act like plain disk for instant restore, making it a great partner for Veeam Backup & Replication. Our testing demonstrated fast Instant VM Recovery, reducing downtime and enabling throughput and response times to be at the levels users need to do business. In addition, ESG Lab validated that the combined solution is fast and easy to set up, and delivers maximum capacity reduction. Veeam's built-in deduplication and compression shrink capacity from the host to speed the backups and save on storage, while ExaGrid deduplication shrinks capacity even further. Moreover, good Veeam practice includes executing synthetic full backups every couple of days, which will be stored using minimal capacity on the ExaGrid target.

Instant restore capabilities are touted by many vendors. ESG's advice is *caveat emptor*, let the buyer beware—just because a vendor says it can instantly restore from deduplicated backup doesn't mean it will actually be instant.

With ExaGrid and Veeam, you can have the benefits of both instant restore and deduplication. Life is full of trade-offs—here's one you don't have to make.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of The Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at 508.482.0188.

⁴ Source: ESG Research Report, [2013 IT Spending Intentions Survey](#), January 2013.